Cost-Effective Data Analytics across Multiple Cloud Regions

Junyi Shu, Xin Jin, Yun Ma, Xuanzhe Liu, Gang Huang

Peking University

Abstract

- Targeting analytics on data stored among **geo-distributed** regions
- Primary goal is to reduce **monetary cost**
- **Prices** of **compute resources** and **data transmission** are taken into account
- Total cost is saved by **15.1%** in our test case

Background

- Business questions sometimes need to be answered based on data from multiple regions

Insights

- Different prices of **compute resources** across regions (0.1713 vs 0.0852)

Region	On-Demand hourly price	Avg hourly price last 3 months	Hourly price as of May 13
us-east-1	0.384	0.1492	0.1546
us-east-2	0.384	0.0821	0.0852
us-west-1	0.448	0.1294	0.1305
us-west-2	0.384	0.1487	0.1526
af-south-1	0.508	0.1436	0.1436
ap-east-1	0.528	0.1506	0.1361
ap-south-1	0.404	0.1119	0.1109
ap-northeast-3	0.496	0.1397	0.1397
ap-northeast-2	0.472	0.1259	0.1259
ap-southeast-1	0.48	0.135	0.133
ap-southeast-2	0.48	0.1447	0.1447
ap-northeast-1	0.496	0.1397	0.1397
ca-central-1	0.428	0.1217	0.1214
eu-central-1	0.46	0.1437	0.1402
eu-west-1	0.428	0.1545	0.1651
eu-west-2	0.444	0.132	0.132
eu-south-1	0.448	0.1357	0.1477
eu-west-3	0.448	0.1273	0.1273
eu-north-1	0.408	0.1147	0.1147
me-south-1	0.471	0.1471	0.1329
sa-east-1	0.612	0.1713	0.1713

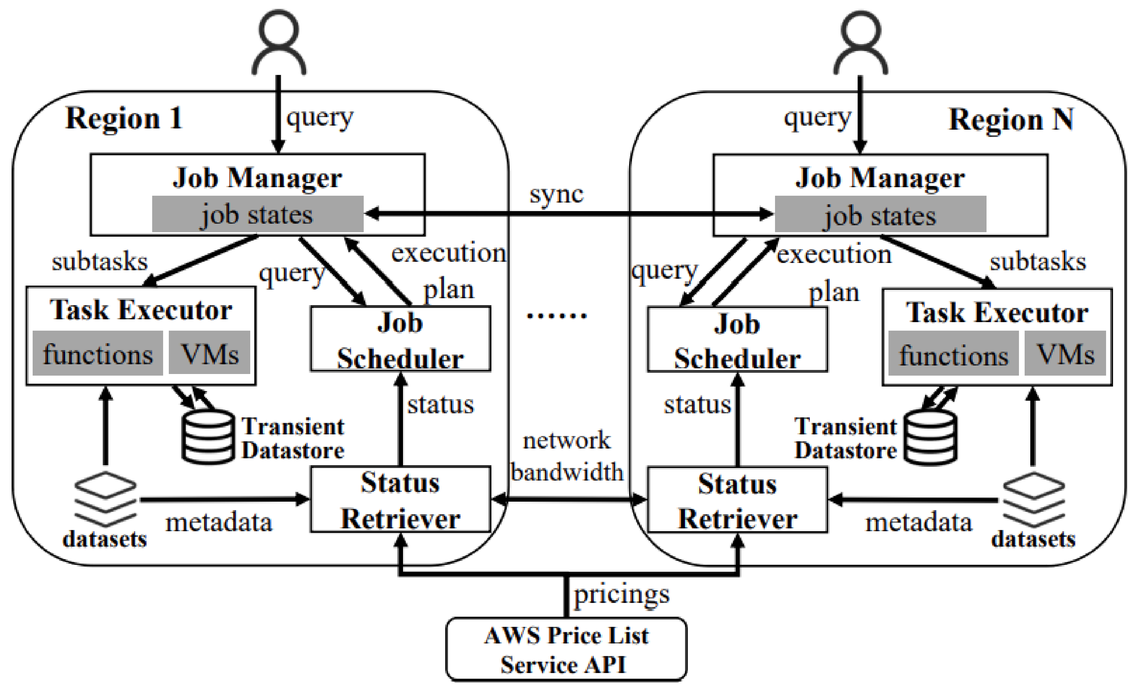
- Different prices of **data transfer** (0.147 vs 0.01) so-called data gravity

	us-east-1	us-east-2	us-west-1	us-west-2	af-south-1	ap-east-1	ap-south-1	ap-northeast-3	ap-northeast-2	ap-southeast-1	ap-southeast-2	ap-northeast-1	ca-central-1	eu-central-1	eu-west-1	eu-west-2	eu-south-1	eu-west-3	eu-north-1	me-south-1	sa-east-1	Internet
us-east-1	0	0.01	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.05
us-east-2	0.01	0	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.05
us-west-1	0.02	0.02	0	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.05
us-west-2	0.02	0.02	0.02	0	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.05
af-south-1	0.147	0.147	0.147	0.147	0	0.147	0.147	0.147	0.147	0.147	0.147	0.147	0.147	0.147	0.147	0.147	0.147	0.147	0.147	0.147	0.147	0.112
ap-east-1	0.09	0.09	0.09	0.09	0.09	0	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.08
ap-south-1	0.086	0.086	0.086	0.086	0.086	0.086	0	0.086	0.086	0.086	0.086	0.086	0.086	0.086	0.086	0.086	0.086	0.086	0.086	0.086	0.086	0.08
ap-northeast-3	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.084
ap-northeast-2	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.108
ap-southeast-1	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.08
ap-southeast-2	0.098	0.098	0.098	0.098	0.098	0.098	0.098	0.098	0.098	0.098	0	0.098	0.098	0.098	0.098	0.098	0.098	0.098	0.098	0.098	0.098	0.092
ap-northeast-1	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.084
ca-central-1	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.05
eu-central-1	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.05
eu-west-1	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0	0.02	0.02	0.02	0.02	0.02	0.02	0.05
eu-west-2	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0	0.02	0.02	0.02	0.02	0.02	0.05
eu-south-1	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0	0.02	0.02	0.02	0.02	0.05
eu-west-3	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0	0.02	0.02	0.02	0.05
eu-north-1	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0	0.02	0.02	0.05
me-south-1	0.1105	0.1105	0.1105	0.1105	0.1105	0.1105	0.1105	0.1105	0.1105	0.1105	0.1105	0.1105	0.1105	0.1105	0.1105	0.1105	0.1105	0.1105	0.1105	0	0.1105	0.065
sa-east-1	0.138	0.138	0.138	0.138	0.138	0.138	0.138	0.138	0.138	0.138	0.138	0.138	0.138	0.138	0.138	0.138	0.138	0.138	0.138	0.138	0	0.114
Internet	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

- Different **bandwidths** between regions (13.22 sec vs 1.08 sec for the same file)

	us-east-1	us-east-2	us-west-1	us-west-2	af-south-1	ap-east-1	ap-south-1	ap-northeast-3	ap-northeast-2	ap-southeast-1	ap-southeast-2	ap-northeast-1	ca-central-1	eu-central-1	eu-west-1	eu-west-2	eu-south-1	eu-west-3	eu-north-1	me-south-1	sa-east-1
us-east-1	1.27	1.18	1.66	1.95	4.55	4.16	3.92	3.38	3.85	4.34	4.15	3.26	1.25	2.24	2.82	2.02	2.35	2.09	2.70	4.20	2.82
us-east-2	1.21	1.42	1.51	3.60	4.71	3.92	4.02	2.98	3.62	4.13	4.16	3.00	1.09	2.37	1.95	2.20	2.54	2.22	2.65	4.38	3.18
us-west-1	1.79	1.48	1.34	1.24	5.54	3.50	4.78	2.52	2.94	3.79	4.56	2.53	2.01	3.36	3.18	3.13	3.42	3.12	3.57	6.11	3.81
us-west-2	1.98	1.47	1.24	1.33	6.27	3.15	4.58	2.32	2.84	3.52	3.08	2.38	1.86	3.27	2.68	2.76	3.23	3.06	3.37	5.43	3.86
af-south-1	4.60	4.65	6.51	5.36	1.08	4.96	3.50	6.92	9.70	4.43	8.65	8.04	4.55	4.09	3.38	3.24	3.88	3.22	3.63	3.19	7.69
ap-east-1	4.61	3.82	3.48	3.34	5.06	1.20	2.34	1.45	1.41	1.56	3.20	1.68	4.09	4.05	4.25	4.06	3.86	4.14	4.31	3.81	6.13
ap-south-1	3.99	4.22	4.95	4.69	3.65	2.46	1.30	2.81	2.87	1.64	3.31	3.25	4.25	2.68	2.73	2.71	2.50	2.53	3.19	1.32	6.05
ap-northeast-3	3.29	2.97	2.64	2.44	9.00	1.55	3.06	1.20	1.24	1.78	2.84	1.25	5.53	4.87	4.29	6.35	4.47	4.85	4.97	5.61	5.38
ap-northeast-2	3.72	3.46	3.04	2.83	9.24	1.33	2.83	1.14	1.35	1.76	3.25	2.51	3.68	4.63	4.66	5.03	4.44	4.88	5.38	3.41	5.80
ap-southeast-1	5.03	4.06	3.92	3.72	4.47	1.51	1.66	1.94	2.00	1.21	2.44	1.92	4.35	3.57	3.83	3.66	3.21	3.67	3.74	2.27	6.48
ap-southeast-2	4.38	3.95	3.06	3.15	10.01	4.70	3.24	2.77	3.13	2.23	1.24	2.73	4.32	5.74	5.41	5.26	4.86	6.55	6.14	3.84	6.56
ap-northeast-1	4.90	3.05	2.59	2.31	8.62	1.60	3.03	1.22	1.90	1.94	2.48	1.36	3.16	4.81	4.11	4.36	4.41	4.67	4.85	3.65	6.17
ca-central-1	1.33	1.21	2.04	1.72	4.54	3.94	4.03	3.11	3.63	4.28	4.08	3.16	1.38	2.32	1.88	2.21	2.49	2.24	2.75	4.28	2.89
eu-central-1	2.27	2.40	3.30	3.15	4.13	3.96	2.66	4.50	4.56	3.32	6.68	4.74	2.35	1.34	1.15	1.15	1.12	1.18	1.16	2.84	4.57
eu-west-1	1.88	1.94	3.05	2.73	3.38	4.59	2.68	4.28	4.73	4.32	5.21	4.25	2.03	1.26	1.40	1.18	1.32	1.28	1.42	3.12	3.87
eu-west-2	1.91	2.07	8.50	2.98	3.12	4.54	2.60	4.56	5.09	4.38	5.29	4.33	2.06	1.18	1.14	1.20	1.13	1.20	1.20	2.84	4.20
eu-south-1	2.39	2.50	3.43	3.58	3.96	3.88	2.57	4.50	4.59	3.19	4.86	4.56	2.41	1.19	1.30	1.21	1.23	1.23	1.22	2.67	4.52
eu-west-3	4.00	2.22	3.23	6.14	3.28	4.39	2.57	4.56	5.00	3.46	7.13	4.58	2.14	1.18	1.14	1.12	1.15	1.21	1.24	2.80	4.44
eu-north-1	2.55	2.58	3.56	3.42	3.70	4.39	3.05	4.89	5.40	3.66	5.87	4.99	2.57	1.19	1.49	1.42	1.18	1.22	1.04	3.08	4.53
me-south-1	4.12	4.36	13.22	5.43	3.22	2.97	1.21	3.32	3.38	2.15	3.76	3.68	4.23	2.84	4.87	2.86	2.63	2.72	3.05	1.22	6.92
sa-east-1	2.65	2.84	3.72	3.89	6.56	7.64	6.96	5.26	5.66	6.39	6.13	5.16	3.07	4.59	3.91	3.93	4.63	4.17	4.46	6.20	1.32

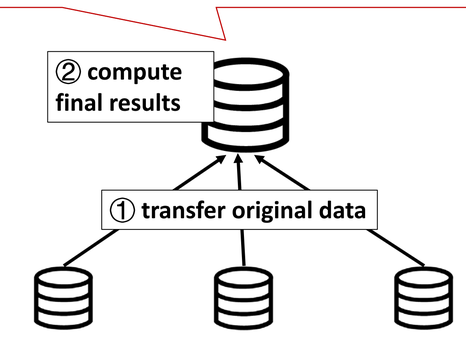
Proposed Architecture



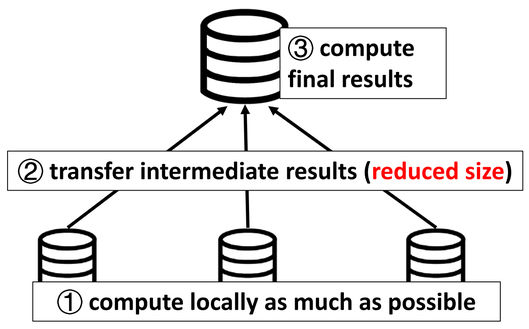
- Continuously **monitor prices and network conditions** to adjust execution plans
- Use **FaaS** rather than provisioning new VMs to reduce resource wasting
- **Decouple data transmission and compute** to mitigate long-waiting caused by slow network

Preliminary Results

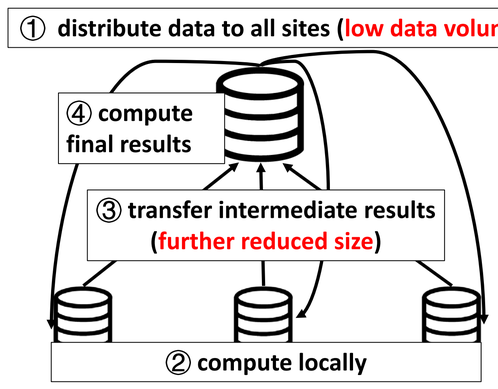
- Aggregation
  - **Has key information to generate final results**
  - **Very inexpensive compute**



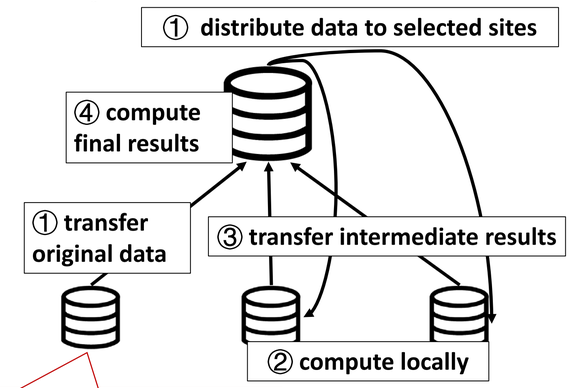
- In-place



- Optimized in-place



- Hybrid



- **Cheap data transfer**
- **Expensive compute**

Strategy	Compute($)	Network($)	Total($)
Aggregation	0.0396	0.2760	0.3155
In-place	0.1007	0.0893	0.1900
Optimized In-place	0.0720	0.0003	0.0723
Hybrid	**0.0580**	**0.0033**	**0.0614**

PEKING UNIVERSITY 北京大学